DOCUMENT RESUME

ED 407 390 SP 037 306

AUTHOR Nasser, Fadia; Glassman, David
TITLE Student Evaluation of University Teaching: Structure and Relationship with Student Characteristics.
PUB DATE Mar 97
NOTE 31p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, March 24-28, 1997).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS College Faculty; College Instruction; Evaluation Criteria; Goodness of Fit; Graduate Students; Higher Education; *Student Attitudes; *Student Characteristics; *Student Evaluation of Teacher Performance; Teacher Effectiveness; *Teacher Student Relationship; Undergraduate Students

ABSTRACT

This paper reports on a study of the relationship between students' characteristics and students' ratings of faculty teaching, using the Faculty Course Evaluation Form (FCEF) at a major southeastern university. In particular, the study investigated: (1) how students rate faculty members on an item-by-item basis (item functioning); (2) what the structure of the FCEF is (test of dimensionality); (3) how students with different characteristics rate faculty members on each of the factors; and (4) which of these factors are potentially problematic in the sense that faculty are rated consistently low on certain factors as opposed to other factors. The FCEF was administered to 3,448 graduate and 2,804 undergraduate students enrolled in 529 classes taught by 260 instructors. The results indicated that among student characteristics, only reasons for taking the course and prior interest in the subject were clearly related to students' ratings at both item and factor levels. Exploratory factor analysis indicated that the FCEF consisted of three major factors and one minor factor. Confirmatory factor analysis showed that the goodness-of-fit of the four factor structure to the data was unsatisfactory. (Contains 33 references.) (ND)

Running head: Students' evaluation of faculty teaching

Student Evaluation of University Teaching: Structure and Relationship with Student Characteristics

Fadia Nasser
David Glassman
The University of Georgia

Paper presented at the Annual Meeting of the American Educational Research Association, March, 24-28, Chicago.

Abstract

The relationships among students' characteristics and students' ratings of faculty teaching were examined using the Faculty Course Evaluation Form (FCEF) at a major southeastern university. The FCEF was administered to 3448 graduate and 2804 undergraduate students enrolled in courses at the College of Education. Included in the study were 529 classes taught by 260 instructors. The results indicated that among student characteristics, only reasons for taking the course and prior interest in the subject were clearly related to students' ratings at both item and factor levels. Exploratory factor analysis indicated that the FCEF consisted of three major factors and one minor factor. Confirmatory factor analysis showed that the goodness-of-fit of the four factor structure to the data was unsatisfactory.

KEY WORD: Teaching evaluation, students' ratings, teaching effectiveness, factor analysis, student characteristics.

Faculty evaluation by students has become an integral part of ɔuntability of education. Over the years, relatively standard ɔocedures for faculty evaluation have evolved, including the four ain types: student, peer, self-and administrative evaluation Mcgee, 1995). One of the most commonly used and still one of the ɔst controversial is student evaluation (Marsh, Overall, & Kesler, 979). This type of evaluation is the focus of the present paper.

Students ratings are used variously to provide the following: . formative feedback to faculty about effectiveness of their eaching, b. a summative measure of teaching effectiveness to be sed in personnel decisions, c. information for students to use in ne selection of instructors and courses, d. an outcome or process escription for research on teaching (Marsh, 1984, 1987, 1989). ile few faculty argue strongly against the usefulness of ratings n providing feedback about instructional effectiveness to the aculty themselves many continue to challenge the use of such atings in personnel decision. Using student evaluation as a measure f teaching effectiveness has also been questioned by many esearchers (e.g., Marsh et al, 1979). Critics of students' ratings rgue that such ratings are biased by variables unrelated to eaching effectiveness. While student ratings are routinely used in any higher education institutions for the first two purposes Newport, 1996), to our knowledge, student ratings are rarely vailable to students and their use for research on teaching is imited.

Both the uses of and the effectiveness of faculty evaluation by tudents are controversial. Student ratings are considered by many eachers to be nothing more than a measure of teacher popularity. ɔme researchers criticize the use of teacher ratings as a tool of anking and/or promoting faculty (Bonetti, 1994). According to these esearchers, students are not qualified to judge whether an nstructor knows the course's subject-material and will not know if ne course is as comprehensive as should be (Lowman, 1984). Other roups of researchers attribute the uselessness of the students atings to poor operational processes used to develop different aculty evaluation instruments which lead to flaws such as unclear tems, or items that do not characterize classroom teaching erformance (Tagomori and Bishop, 1995).

Proponents of students' evaluation of faculty teaching argue hat as an appropriately designed survey instrument, student valuations are valuable, reliable and valid (e.g., see Cohen 1981, ckeachie 1986; Marsh 1984, Marsh and Ware 1982; Murray 1983; Seldin 984). According to this camp of researchers, college student are professional teacher watchers" and, if asked relevant questions hat are within their experiential background, can make fair and ound judgements about teaching (Miller, 1988).

A major reason for the complexity of evaluating teaching in higher education is the great difficulty of defining effective teaching and the lack of agreement about what "good teaching" is (Goodwin & Stephens, 1993). Items on many of the evaluation instruments are considered reflections and measures of effective teaching as viewed by students, faculty, and the instrument's designer/s. These items are what teachers, students and other educational professional collectively specify as behaviors that constitute effective teaching. Critics of such instruments questioned their validity and whether they truly reflect teaching effectiveness because such definition of effective teaching is not tied to student outcomes (Tuckman, 1995).

Bonittee (1994) distinguished between two types of evaluation questionnaires which are conducted for information and questionnaires conducted for action. Those conducted for information tend to consist of a list of specific technical questions about the structure of the course, the structure of the lectures, and the clarity, enthusiasm, audibility and motivational ability of the lecturer. The intention of such questions is diagnostic, to provide a flow of information to instructors on the quality and character of their performance, leaving it to individuals to remedy any defects identified.

At the other extreme are questionnaires conducted for action. The range of actions which can be informed by student questionnaires is broad. They cover the possibility of changes in the course content, course difficulty, teaching methods and prescribed textbooks. More acutely they raise the possibility of using questionnaire results for managerial actions like tenure awards, allocation of staff among courses, and recommendations for the award of performance-related pay supplements. A further possible use is as means for institution-wide resources allocation.

Dimensionality of Teaching Evaluation

Effective teaching is a multidimensional construct. Thus it is not surprising that a large body of research has shown that students' evaluation of teaching effectiveness designed to reflect effective teaching are also multidimensional (Marsh, 1987, 1991). Some researchers argue that students' evaluations of teaching effectiveness are best considered as a relatively unidimensional construct, whereas others argue for multidimensional perspective (Marsh, 1991). Others proposed a comprise in that "effective teaching may be described as unitarily and multidimensionally in a way analogous to the way Weschler's tests operationally define intelligence in both general and specific terms" (Abrami, 1985, p. 214). For personnel decisions, some researchers argue that a single score is more useful than multidimensional ratings (Abrami, 1988, 1989), whereas others argue the opposite (Marsh, 1987). Marsh (1989) noted that for the three uses of students' evaluations listed earlier, there appears to be a general agreement that appropriately constructed multidimensions are more useful than a single summary score.

The literature on students' evaluation of teaching effectiveness contains several examples of well-constructed instruments with clearly defined factor structures that provide measures of distinct components of teaching effectiveness (for a list of these instruments see Marsh, 1991). Commenting on these instruments, Marsh (1987) noted that the systematic approach used in the development of these instruments and the similarity of the factors that they measure support their construct validity. Factor analyses of responses to each of these instruments have provided clear support for the structure they were designed to measure, demonstrating that the students' evaluations measure distinct dimensions of teaching effectiveness. Several researchers also tested higher order structures of students' evaluations of teaching effectiveness. Feldman (1976) proposed a model with three higher-order factors which he labeled presentation, facilitation and regulations. These three categories are first-order factors because each of his categories consisted of one item. Frey (1978) proposed two higher-order factors to seven first-order factors of his 21-item Endeavor instrument, and argued for the usefulness of the two global factors that he called pedagogical skill and rapport.

Marsh (1991) pointed out that the higher order structure described by Frey (1976) is actually first order structure, because it is based on one item from each category. He also noted that Frey's factor structure based only on exploratory factor analysis and the fitness of that model has never been tested. Marsh (1991) employed confirmatory factor analysis to test the goodness of fit of four a priori higher-order factor structure (1-4 second order-factor models) of his nine first-order factor Students' Evaluation of Educational Quality (SEEQ) questionnaire. Marsh's (1991) results indicated that only the four second-order factor model fit the data better and explained more variance in the first-order factors than did the models posting fewer higher-order factors. These results provide a strong support for the claim that students' evaluations of teaching effectiveness are multidimensional and that their responses cannot be adequately explained by one or even a small number of factors.

Student Characteristics and Student Evaluations

Another reason for the complexity of evaluations of teaching effectiveness is rooted in the argument that such evaluations are biased by variables unrelated to teaching effectiveness. Some student's characteristics, course characteristics and teacher characteristics have been discussed in the literature as being responsible for biased students' evaluations of faculty teaching. However, Marsh (1987) discussed this argument about bias in students' evaluations and concluded that it often stems from misuse and misunderstanding of the concept of bias. In Marsh's view, the differences in students' evaluations do not always indicate bias but true differences in evaluation between groups of students with different characteristics, or classes with different characteristics

because in both cases these variables are related to teaching effectiveness (for detailed discussion of bias see Marsh, 1987).

Hundreds of studies have used a variety of approaches to examine the influence of many background characteristics on students' evaluations of teaching effectiveness, and a comprehensive review is beyond the scope of this study. Empirical findings in this area have been reviewed by many researchers (e.g, Centra, 1979; Feldman, 1976 1983, 1984; Marsh, 1983, 1984).

According to Marsh, (1987), over 50% of the faculty who were asked which of a list of 17 characteristics would cause bias to student ratings cited the following: course difficulty, grading leniency, instructor popularity, student interest in the subject before taking the course, course work load, class size, reason for taking the course, and student GPA. Marsh (1978) examined the relations among a wide variety of background characteristics, but concluded that most of the variance in student's evaluations that could be accounted for by the entire set could be explained by class size; workload/difficulty; prior subject interest; expected grades; and the reason for taking a course. However, there is considerable evidence that most background variables such as class size, reason for taking the course, workload, and grade point average have little relationship to student ratings of faculty teaching (Marsh, 1978).

Of interest to the present study is the relationship between student evaluations and the following student characteristic: level of education, reason for taking the class, GPA, percentage of class meetings attended, hours per week devoted to the course outside the class, and interest prior to taking the course. The direction and the magnitude of the relationship between the aforementioned student characteristics and student ratings differ across studies. These differences can be attributed, in part, to different methods employed for analyzing the data, different questionnaires, and different institutions. Based on his own studies and on reviews made by other researchers, Marsh (1987) pointed out that for most of the relationship between students characteristics and student evaluations of effective teaching, the effects tend to be small, and the directions of the effects are sometimes inconsistent. Marsh's claim that a variety of variables that could potentially influence student evaluations apparently have little effect fortified similar conclusions drawn earlier by Centra (1979), Menges (1973) and others.

Despite the inconsistent findings and the small effect of student characteristics on their evaluation of teaching effectiveness there is considerable evidence that some of these characteristics, such as level of education (graduates vs. undergraduate), reason for taking the class, prior interest, and workload, are positively related to student evaluations (Marsh, 1987). Other characteristics such as GPA had negligible influence on student ratings while percentage of attendance, to our knowledge, has not been examined in previous research.

Purpose of the Study

The purposes of this study were to investigate (1) how students rate faculty members on the Faculty Course Evaluation Form (FCEF) on an item by item basis (item functioning), (2) what the structure of the FCEF is (test of dimensionality, (3) how students with different characteristics rate faculty members on each of the factors, (4) and which of these factors is/are potentially problematic in the sense that faculty are rated consistently low on certain factors as opposed to other factors.

Method

Sample

The sample consists of 6252 graduate (3448) and undergraduate (2804) students enrolled in Fall 95 and Winter 96 in the College of Education at a major southeastern university. Included in these data were 521 separate classes, taught by 260 instructors. Data will be analyzed utilizing each instructor as the unit of analysis.

Instrument

The Faculty-Course Evaluation Form (FCEF) was first developed in 1972. Thirty-eight items representing the primary dimensions of teacher performance reported by Deshpande, Webb, and Marks (1970) were selected. A Likert-type rating scale was applied to 36 of the items as a frequency indicator. The last two items were open-ended summarizations of the course and instructor. The original instrument also included five items relating to student characteristics.

Approximately 5,000 undergraduate and graduate students were utilized for the refinement of the instrument. These students rated a total of 222 instructors. The individual student was the unit of analysis. Factor analysis yielded five factors (Subject Organization and Competence, Motivation-Stimulation, Instructor-Student Relations, Reasonable Work Load and Tests, and Clearness of Grading Procedures) which were then weighted by a survey of faculty members. The resulting factors and their corresponding weights were as follows: Subject Organization-35, Motivation-Stimulation-30, Instructor Relations-16, Work Load-10, and Grading-9. These factors were moderately correlated with one another. The total score for the instrument was the sum of the weighted averages of each scale.

As a result of this preliminary analysis, the original instrument was refined. Ten items were discarded; three student characteristic items were added resulting in a total of eight. The two opened-ended items were converted to the same Likert-type scale. The resulting 36 items comprises the current FCEF.- eight student characteristics; 26 specific and two overall items (see Appendix A). The weighting scale is not utilized in the computation of the total scores.

An examination of this instrument was performed in the 1984-85 academic year. A principle component analysis with orthogonal rotation procedure was employed using both individual students (N=1346) and class means (N=97) as the units of analysis. Both levels revealed remarkably similar patterns of factor loadings. Four factors were extracted (Motivation/Stimulation, Subject Matter & Organization, Testing/Grading Practices, and Workload).The reliablilities of the resulting factors ranged from .86-.96.

Procedure

The FCEF is administered to each class at the end of each academic quarter. The responses are scannned and descriptive results and summary of students' comments are reported to instructors and heads of their departments. Results of these evaluations are used for personnel decisions and as feedback for instructors.

Data Analysis

Since individual observations within classrooms are more likely to be dependent, and evaluations of the same instructor within different classrooms are also likely to be dependent, data were aggregated to instructor level.

The relationships between student characteristics (level of education, reason for taking the course, GPA, percentage of classes attended, hours per week devoted to the course outside the class, interest in subject prior to taking the course) and student evaluations were examined using t-test and ANOVA. This was done on a item-by-item and per factor basis,

In all these analyses item means for student characteristic subgroups were calculated per instructor and used. For example, mean scores for undergraduates and for graduate students were calculated on each item per instructor and the differences between these means were examined by t-test.

Common factor analysis was conducted to explore the structure underlying the FCEF. Confirmatory factor analysis was conducted to test the goodness-of-fit of the resulting model to the data. Factor scores for the resulting factors were calculated by summing the scores of the items which are loaded on each factor and the relationships between these scores and student characteristics subgroups were examined. All types of analysis were performed using version 6.10 of SAS for PC.

Results

Descriptive Statistics

Table 1 summarizes the means and standard deviations of items 9-36 based on the instructor as the unit of analysis.

Insert Table 1 here.

All item means shown in Table 1 are relatively high. Only items 31 and 35 were slightly smaller than 4.0 (on a scale of 5.0). The values of the standard deviations indicate that the variability of the rating was not high. The reliability of the instrument measured by Chrombach alpha was .97.

Student Characteristics and Ratings

Item Level

The relationships between each of the student characteristic items 1,2,3,6,7 and 8 with items 9-36 were examined using t-test and ANOVA. The results of the item level analysis are presented for each of these items.

Item1-Education Level

Item1 was dichotomized into undergraduate/graduate levels. Only 26 out of 260 instructors taught both levels and were included in the analysis.

Insert Table 2 here

While the mean scores of graduate students were higher than those of undergraduates on 27 of the 28 items, the results of the t-test shown in Table 2 indicated that graduate students rated instructors significantly higher than undergraduate students on only four items (13,22,26, and 31). However, these items have, seemingly, nothing in common.

Item2-Reason for Taking the Course

Item2 consisted of five possible choices. Inspection of the data indicated that only a few students selected the class because they thought they could make a good grade; therefore this choice was not considered in the analysis of this item. Even though the overall difference between the ratings of the groups was not of primary interest in the analysis of this item, it is informative to mention that 17 of the 28 ANOVA tests were statistically significant ($p<.05$).

Insert Table 3 here

Three contrasts were of particular interest to this study, and the results involving these contrasts are summarized in Table 3. The first contrast compared the rating of students who selected the course because of their interest in the subject to students who were required to take the course. The former group rated instructors significantly higher than the later group on all items except items 12,16,18,25,26,27,31,and 32.

The second contrast compared the rating of students enrolled in the course because of the recommendation of their advisor to students who enrolled because they were required to take the course.

The first group rated instructors significantly higher than the second group only on items 10,15,17,20,and 24. Inspection of the content of these items indicates that they pertain to the ability of the instructor to motivate or stimulate students.

The third contrast compared the rating of students who enrolled in the course because of either the reputation of the instructor or because of their interest in the subject to students who enrolled in the course because they were required to take the course or because of their advisors recommendations. The first combined group rated instructors significantly higher than the second combined group on nearly half of the items (9,10,14,15,17,20,21,24,29,30,35, and 36).

Item3-Grade Point Average

Item3 included few students with a GPA less than 2.0; therefore this group was not included in the analysis of this item. Student grade point average had little influence on faculty ratings. Out of the 28 items examined only the ratings on item13 ($F=4.05$, $df=3$, $p=.008$) and item25 ($F=2.86$, $df=3$, $p=.038$) were significantly different across students with different GPA's (omnibus tests). Follow-up dependent t-tests on those two items indicated that students with GPA of 3.5-4.0 rated instructors significantly higher than students with GPA of 2.0-2.49 ($t=.2.81$, $n=57$, $p=.007$; $t=2.88$, $n=58$, $p=.006$ for items 13 and 25 respectively. In addition, students with 3.0-3.49 rated instructors higher than students with a GPA of 2.0-2.49 on item13 ($t=3.79$, $n=58$, $p=001$).

Item6-Percentage of Class Meetings Attended

Ninety-five percent of all student responses were in the last two categories (60-80%, 80-100%). As a result only these two groups of students were considered in the analysis of this item.

Insert Table 4 here

As shown in Table 4 the mean ratings of students who attended 80-100% of classes were higher than the mean ratings of students who attended 60-80% of the classes 27 of the 28 items. However, the differences were statistically significant only on items 10,15,18,26,30, and 31.

It was interesting to note that the ratings on item26 were significant, but not on item32 ($p=.029$, $p=.277$ for items26 and item32 respectively). These two items are almost identical in content, yet yielded different results.

Item7-Hours per Week Devoted to Course Outside Classroom

All five possible responses were considered in the analysis of this item. In general, students who devoted the most and the least hours rated instructors lower than students in the middle categories. Noticeable differences were not observed among the ratings of students in the three middle categories. T-test results

indicated significant differences in student ratings on only three items (19,32, and 35). The direction of these differences was not uniform, and these items have little in common.

Item8-Interest in Subject Prior to Taking the Course

All five responses were considered in the analysis of this item. In general, level of interest in the subject was positively related to students' ratings of instructors.

Insert Table 5 here

Table 5 summarizes the major differences indicated by the analysis of this item. The mean scores of students who had very great interest in the subject were higher than those of students who had average or small interest in the subject on all items except item9. This clear pattern of differences was not observed when student with average and small interest in the subject were compared. T-test analyses indicated that in almost all items the differences in ratings between students with very great interest and students with average or small interest were statistically significant.

Student Characteristics and Ratings on Overall Items

Items 35 and 36 are overall evaluations of course and instructor. As such, the implication of the results involving these items should be considered differently from items 9-34. There was a significant difference in the ratings on items 35 and 36 with respect to student characteristic items 2 and 8. Students who took the course because of their interest in the subject rated the course and the instructor significantly higher than those who were required to take it. Students who took the course because of their interest or instructors reputation also rated the course and instructor higher than students who were required to take the course or because it was recommended. Students who took the course because it was recommended rated instructors higher than students who were required to take the course. As the level of interest prior to taking the course increased, course and instructor ratings also increased. This was particularly true for course ratings.

FCEF Dimensionality

Principle axis factor extraction with an oblique rotation was performed on item9 through item34. Item35 and item36 were excluded from the factor analysis because they are overall course and instructor evaluation items.

Insert Table 6 here

The factor analysis yielded three major factors along with one minor factor which consisted of only two items. As can be seen in

Table 6 items 9,10,13,14,15,19,21,24,28,30,31, and 34 loaded on the first factor. These items involved the instructor's ability to motivate and stimulate the students; therefore this factor was named Motivation/Stimulation. Factor two contained item 12,16,20,23,27,29, and 33. These items involved instructors subject knowledge and organizational skills, hence it was named Subject Matter/Organization. Factor three included items 11,17,18,22, and 25. These items involved testing and grading procedures of the instructor; therefore it was named Testing/Grading Practices. The forth factor consisted of items 26 and 32. These items relate to the workload assigned by the instructor. Items 17 and 20 had complex loadings, that is, item17 loaded equally on factors two and three while item20 loaded equally on factors 1 and 2.

Because it is debatable whether two items could constitute a factor, a three factor solution was considered. However, within the three factor solution, these two items clustered together, once again, into a factor while the other three factors collapsed into two uninterpretable factors. As a result it was decided to maintain the four factor solution which accounted for 79% of the variance in the data. The internal consistency as measured by Chrombach's alpha was .77,.62,.68,.79 for factors 1-4 respectively. The inter-factor correlations among the factors ranged from .47-.86. The first factor included the largest number of items (12) but not the highest internal consistency. Inspection of the content of these two items revealed that they are almost identical.

Factor scores were obtained to determine whether ratings were uniform across factors. This was done by computing the mean score of each factor based on the raw scores on each of the items loaded in that factor. The mean scores for the four factors were 4.32, 4.45, 4.21, 4.02 for factors 1-4 respectively. These results indicate that first two factors received the highest ratings. While the Workload factor received the lowest mean ratings.

Confirmatory factor analysis was performed on the four factor solution assuming simple structure to test the goodness-of-fit of this model to the data. The results of these analysis indicated that the goodness-of-fit of the four factor model was far from being satisfactory (GFI=.061, RMR=.020, chi-square=1496.96;chi-square/df=5.1; CFA=.809; TLI=.770)[1].

When the two complex items were removed from the solution, one at a time, the goodness-of-fit improved slightly but it remained far from satisfactory. Also, a model including only the three major factors was examined and a slight improvement in the goodness-of-fit over the four factor solution was obtained but the three factor solution remained unsatisfactory. Inspection of alternative models was beyond the scope of the study.

---

[1] GFI=Goodness-of-Fit Index; RMR=Root Mean Square Residual; CFI=Bentler's Comparitive Index; TLI=Tucker-Lewis Index

Student Characteristics and Ratings-Factor Level

The relationship between student characteristics (items 1,2,3,6,7,and 8) and their rating on each of the four factors was examined and the major results are summarized in Table 7.

Insert Table 7 here

As in the case of individual items, only instructors who had students responding to, at least, two categories in the student characteristic items were included in this analysis.

As indicated in Table 7, overall, graduate students, students who chose the class because of their interest in the subject or instructor's reputation, students who maintained a GPA of 3.5-4.0, students who attended 80-100% of class meetings, and students who had very great interest in the subject rated instructors higher than students in other categories on factors 1,2, and 4. On factor 3 the trend of ratings across student characteristics was mixed.

In terms of statistical significance, graduate students rated instructors higher than undergraduates only on factor 3 (Grading/Testing). The combined group of students who took the course because of interest in the subject or instructor's reputation, rated the instructors higher than the combined group of students who were required to take the course or took the class because of their advisor's recommendation on factors 1 and 2. Students who took the course because of the instructors reputation rated instructor higher than those who took the class because it was recommended by their advisor only on factor 1. In addition students who took the class because of their interest in the subject rated instructor higher than students who took the class because it was required.

No significant differences were observed in the ratings of students with different GPA's on any of the four factors. Students who attended 80-100% of class meetings rated instructors higher than those who attended only 60-80% on factor 1 and factor 2. The results concerning item7 indicated that students who devoted the most and the least number of hours rated instructors lower than students in the middle categories on factors 1,2, and 3 but not on factor four. This was consistent with the item level findings. This trend for the first three factors was not statistically significant, with the exception of one comparison between subgroups A and C (see Table 7). With regards to factor four, however, the trend was different. As the number of hours devoted to the course outside the classroom increased, the ratings decreased, except for students who devoted 6-8 hours (middle category) to the course outside the classroom. Three comparisons on this factor were statistically significant. Ratings of students who devoted more than 12 hours to the course were higher

than those of students who devoted 0-2 or 3-5 hours. Also, students who devoted 3-5 hours rated instructor higher than students who devoted 6-8 hours factor 4.

Consistent with the item level results, levels of interest in the subject prior to taking the course was clearly related to student ratings. Except for students who had nil interest in the subject, the ratings on factors 1 and 2 increased as a function of the level of interest in the subject. Ratings on factors 3 and 4 increased as a function of level of interest including students who had nil interest in the subject. Six pairwise comparisons on factor 1, four pairwise comparisons on factors 2 and 4, and two pairwise comparisons on factor 3 were statistically significant (see Table 7).

Discussion

This study was designed to answer four basic questions. First, how do students rate faculty members on the FCEF on an item by item basis ? Based on a limited number of instructors (26 out of 260), no differences were observed in the rating of graduate vs. undergraduate students. This finding is consistent with that of Menges (1973), Centra (1979), and Marsh (1987). Although mean scores for graduate students were consistently higher than those of undergraduates on nearly all items, few were statistically significant.

In general, it was found that the reason for taking the course was an influential variable in the determination of instructor ratings. As expected, students who selected the course because they were interested in the subject or because of instructor reputation rated instructors higher than students who were required or advised to take the course. One interpretation of these results is that higher interest in the subject creates a more favorable learning environment and facilitates effective teaching, and this effect is reflected in the student ratings (Marsh,1987).

Consistent with previous research, the results of this study indicated that GPA had only a minute effect on student ratings. In other words, students with higher GPA rate instructors about the same on almost all items. This is in contrast to the faculty view of characteristics that cause bias in student ratings as described by Marsh (1987).

Although only two categories of class attendance were included, in the examination of item6, the results implied that the percentage of class meetings attended had some influence on the ratings of instructors. Ratings were increased as a function of percentage of classes attended. The attendance influence on ratings may be related to the interest in the subject and reason for taking the course. In both cases it can be argued that interested students will attend more meetings than uninterested students; hence, they rate instructors higher.

The pattern of the relationship between number of hours per week devoted to the course and student ratings is interesting. In general students who devoted the lowest and the highest number of hours to the course rated instructors lower than students who devoted number of hours in between these extremes. However, most of these differences were not statistically significant. These results can be explained by the fact that students who devoted few hours may have less interest in the subject; therefore they rated instructors lower than others. On the other extreme, students who devoted 12 hours or more may develop a negative attitude which is reflected in the low ratings. The interpretation of these results is based on speculation because, to our knowledge, there no existing body of literature pertaining to this issue.

Student level of interest in the subject prior to taking the course seems to be more related to students ratings than all other student characteristics examined in this study. The interpretation of the relationship of interest in the subject and ratings is similar to the interpretation of the relationship between reasons for taking the course and ratings. This finding is consistent with those of Marsh (1980, 1983).

The findings concerning items 35 and 36 are consistent with those of items 9 to 34. In other words, reasons for taking the course and prior interest were also the most influential student characteristic on these two items.

The second research question in this study pertained to the structure of the FCEF. The four-factor structure yielded in the exploratory factor analysis is similar to the results of principle component analysis reported by Payne (1985). Of the resulting four factors, one factor included 12 items, while another included only two redundant items. Compared to evaluation instruments discussed in the literature such as the SEEQ (Marsh, 1991), the FCEF included only a few dimensions of teaching effectiveness. Another issue worth mentioning involves the consideration of the workload factor as a representation of teaching effectiveness. In previous research concerning student ratings, workload was treated as a background variable rather than a factor of teaching effectiveness (Marsh, 1987).

The results of the confirmatory factor analysis which indicated unsatisfactory goodness-of-fit of the four factor solution of the data should be treated cautiously. Because alternative models were not examined, and further research is needed to establish or refute the four factor structure.

The third question involved the relationship between student characteristics and ratings on each of the four factors. In general the relationship between student characteristics and ratings on the factors are consistent with their ratings on the individual items, however, clear patterns of relationships were observed with the factor levels. Reason for taking the course and interest in the

course seem to be the most influential student characteristics also at the factor level. The ratings increased as a function of the level of interest of the subject. This was true for all categories of item8 except for students who had nil interest in the subject. The ratings of this subgroup of students on factor 1 and 2 were not the lowest as they were for factors three and four. One interpretation of these results is that students in this category were less critical of instructors motivational or organizational abilities and were more critical when it came to grading and workload.

Concerning the forth question which pertained to possible differences in factor means, there was no indication that any of the four factors' means differed substantially. However, factor 4 received the lowest mean score and had the highest reliability. The low mean score could be the result of the negatively phrased items included in this factor. The high reliability of factor 4 resulted from the high reliability of the two items in this factor (alpha=1.00 and alpha=.75 for items 26 and 32 respectively).

Recommendations

The results of these analyses suggest there are a number of issues to be addressed pertaining to the overall usefulness of the FCEF. The most obvious examples of these issues are addressed here.

First, several of the response choices available for selection on the student characteristic items (1-8) were not selected with sufficient frequency to warrant their inclusion. Examples of these response choices include "thought I could make a good grade" on item2; "less than 2.0" on item3; and "0-20", "20-40", and "40-60" on item6. This was particularly evident with respect to item6 (percentage of class meetings attended). Since over 95% of all students who responded to this item chose "60-80" or "80-100", it may be desirable to decrease the variability of those two choices (i.e. less than 60, 60-70, 70-80, 80-90, 90-100).

Although factor analysis extraction yielded a four factor solution to this instrument, this solution was problematic. The first problem involves factor 1 (Motivation/Stimulation). Twelve items which comprised 46% of the items loaded on factor one, while only 14 loaded on the remaining three factors. Considering the internal consistency of factor 1 was only .77, further examination of its items is recommended. The elimination or rewriting of some of these items may be required.

The second problem concerns Factor 4. This factor not only contains two redundant items, but it could be argued that the workload construct is not a valid measure or even a dimension of "teaching effectiveness".

When considering whether the FCEF or any other evaluative instrument is valid, one must contemplate the intent for which it was developed. If the FCEF was designed to evaluate the teaching

effectiveness of instructors, the constructs should reflect those issues deemed important to the institution or relevant stakeholders. This should be the first consideration in its development. Next, a sufficient number of items per construct should be developed and piloted. Whether the four constructs contained in this instrument are valid indicators of teaching effectiveness will not be determined here, but this issue should be seriously considered if this instrument is to be refined.

While the FCEF was developed as a tool to solicit feedback for instructors and administrators, it may also be of great value to students. Students have assorted priorities and concerns when selecting a program of study. Information pertaining to instructor abilities and teaching style may allow students to make more informative and therefore competent decisions when selecting classes.

In summary, it is clear that the validity, reliablity and usefulness of student evaluations will remain a controversial topic in higher education. However,an ageed upon definition of teaching effectiveness along with clear purpose of the evaluation can assist in developing better measures of effective teaching. There is much to do in terms of research in order to establish the validity of teaching evaluation.

References

Abrami, P. C. (1985). Dimensions of effective college instruction. Review of Higher Education, 8, 211-228.

Abrami, P. C. (1988). SEEQ and ye shall find: a review of Marsh's students' Evaluations of university Teaching. Instructional Evaluation, 9, 19-27.

Abrami, P. C. (1989). SEEQing the truth about student ratings of instruction. Educational Research, 81, 43-45.

Bonetti, S. (1994). On the use of student questionnaires. Higher Education Review, 26, 57-64.

Cohen, P. A. (1981). Student ratings of instruction and student achievement: A meta-analysis of multisection validity studies. Review of Educational Research, 51, 31-49.

Centra, J. A. (1979). Determining faculty effectiveness. San Francisco: Jossey-Bass.

Deshpande, A.S.; Webb, S.C.; and Marks, E.(1970). Student perceptions of engineering instructor behaviors and their relationships to the evaluation of instructors and courses. American Educational Research Journal 7. 289-305.

Frey, P. W. (1978). A two dimensional analysis of student rating of instruction. Research in Higher Education, 9, 69-91.

Feldman, K. A. (1976). A superior college teacher from the student's view. Research in Higher Education, 5, 243-288.

Feldman, K. A. (1983). The seniority and instructional experience of students. Research in Higher Education, 18, 3-124.

Feldman, K. A. (1984). Class d students' evaluations of college teacher and course: A closer look. Research in Higher Education, 21, 45-116.

Goodwin, L. D., & Stevens, E. A. (1993). The influence of gender on university faculty members' perceptions of 'good' teaching. Journal of Higher Education, 64, 166-165.

Lowman, J. (1984). Mastering the techniques of teaching. San Francisco: Jossey-Bass.

Marsh, H. W. (1978). Students' evaluations of instructional effectiveness: relationship to student, course and instructor characteristics. Paper presented at the Annual Meeting of the American Educational Research Association, March 27-31, Toronto, Canada.

Marsh, H. W. (1980). Prior subject interest, students' evaluations, and instructional effectiveness. Paper Presented at the Annual Meeting of the American Educational Research Association, April 7-11, Boston Ma.

Marsh, H. W. (1983). Multidimensional ratings of teaching effectiveness by students from different academic settings and their relation to student/course/instructor charactersitics. Journal of Educational Psychology, 75, 150-166.

Marsh, H. W. (1984). Student's evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and utility. Journal of Educational Psychology, 76, 707-754.

Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research [Entire issue]. International Journal of Educational Research, 11(3).

Marsh, H. W. (1989). Responses to review of "students' evaluations of university teaching: Research findings, methodological issues, and directions for future research". Instructional Evaluation, 10, 5-9.

Marsh, H. W. (1991). Multidimensional students' evaluations of teaching effectiveness: A test of alternative higher-order structures. Journal of Educational Psychology, 83, 285-296.

Marsh, H. W., & Ware, J. E. (1982). Effects of expressiveness content coverage, and incentive on multidimensional student rating scales: New interpretations of Dr. Fox effect. Journal of Educational Psychology, 74, 126-134.

Marsh, H. W., Overall, U. J., & Kesler, P. S. (1979). Validity of student evaluation of instructional effectiveness: A comparison of faculty self-evaluations and evaluation by their students. Journal of Educational Psychology, 71, 149-160.

Mcgee, R. (1995). Faculty evaluation procedures in 11 western community colleges. Community College Journal of Research and Practice, 19, 341-348.

McKeachie, W. J. (1986). Teaching Tips (8th ed.). Lexington, MA: D. C. Heath and Co.

Menges, R. J. (1973). The new reporters: students rate instruction. New Directions For Higher Education, 1, 59-76.

Miller, R. I. (1988). Evaluating faculty for promotion and tenure. San Francisco: Jossey Bass.

Murray, H. G. (1983). Low inference classroom teaching behaviors and student rating of college effectiveness. Journal of Educational Psychology, 75, 138-149.

Newport, F. J. (1996). Rating teaching in the USA: probing the qualifications of student raters and novice teachers. Assessment & Evaluation in Higher Education, 21, 17-21.

Payne, D.A. (1974). The Assessment of Learning; Cognitive and Affective. Lexington, Ma: D.C. Heath and Company.

Payne, D.A. (1985). Research on the University of Georgia Faculty-Course Evaluation Form (FCEF). (Tech. Report) Athens, Georgia: University of Georgia.

Seldin, P. (1984). Changing practices in faculty education. San Francisco: Jossey Bass.

Tagomori, H. T., & Bishop, L. A. (1995). Student evaluation of teaching: Flaws in the instruments. Thought & Action, 11, 63-78.

Tuckman, W. B. (1995). assessing effective teaching. Peabody Journal of Education, 20, 127-138.

Appendix A

FACULTY - COURSE EVALUATION FORM

PURPOSES AND USES

The College of Education is interested in improving its instructional programs. This questionnaire gives you an opportunity to express anonymously your view of this course and the way it has been taught. The purpose of obtaining the information is two-fold: to assist in diagnostic or self-improvement type decisions and to assist as one criterion in administrative decisions.

Tabulations of your answers will be given (1) to your professor so that he/she can study them and use your collective responses to improve his/her performance in class, and (2) to the Department head so that he/she can use them as one criterion for a faculty member's annual evaluation. These evaluations will be made available midway or later in the following quarter. The information you provide will be kept anonymous. For this reason you should NOT place your name on this form.

Please follow the instructions carefully.

1. What is your class standing ?
   O Fresh. O Soph. O Jr. O Sr. O Grad.
2. Which one of the following was your most important reason for selecting this course ?
   O It was required O Teacher's excellent reputation
   O Advisor's recommendation O Thought I could make a good grade
   O Subject was of interest
3. What is your present grade point average ?
   (Leave blank if not yet established)
   O less than 2.0 O 2.5 - 2.99 O 3.5-4.0
   O 2.0 - 2.49 O 3.0 - 3.49
4. What grade do you expect to get in this course ?
   O A O B O C O D O E
5. What do you feel you deserve ?
   O A O B O C O D O E
6. What percentage of the class meetings did you attend ?
   O 0 - 20 O 20 - 40 O 40 - 60 O 60 - 80 O 80 - 100
7. How many hour per week did you devote to this course outside of class ?
   O 0 - 2 O 3 - 5 O 6 - 8 O 9 - 11
   O 12 or more
8. What was your interest in this subject prior to taking this course ?
   O Nil O Average O Very Great
   O Small O Substantial

ON THE FOLLOWING ITEMS, ESTIMATE HOW FREQUENTLY YOU FEEL THE FOLLOWING OCCURRED.

1 Almost never 2 Infrequently 3 Occasionally 4 Often 5 Almost Always

---

9. The instructor was willing to give individual assistance outside of class.

10. The instructor encouraged students to think for themselves.

11. The instructor gave tests that were reasonable in length.

12. The instructor spent time on unimportant and irrelevant materials.

13. The instructor pitched the presentation above the heads of the students.

14. The instructor encouraged the students to ask questions.

15. The instructor tried to get you to see beyond the limits of this course.

16. The instructor was well prepared each day.

17. The instructor clearly described the grading procedures.

18. Test content was representative of assigned material.

19. The instructor stimulated the intellectual curiosity of the students.

20. The instructor was enthusiastic about the subject.

21. The instructor was clear about basic principles.

22. The instructor clearly indicated what materials tests would cover.

23. The instructor kept the course moving at a steady pace.

24. The instructor tried to stimulate creative abilities.

25. The instructor gave advice on how to study for the course.

26. The instructor assigned a lot of burdensome busy work.

27. The instructor gave presentations that were logically arranged.

28. The instructor tried to increase the interests of class members in the subject.

29. The instructors information seemed up-to-date.

30. In this class I felt free to express my opinions.

31. The instructor explained text materials that were confusing to students.

32. The instructor demanded an unreasonably large amount of work.

33. The instructor seemed well informed about the material presented.

34. The instructor recognized student's difficulties in understanding new material.

---

35. How would you rate the over-all value of this course ?

O Poor O Very Good
O Fair O Superior
O Good

36. How would you rate the teaching ability of this instructor ?

O Poor O Fair O Good O Very Good O Superior

Table 1
Means and Standard Deviations of Items 9-36

Item	N	Mean	SD
9	260	4.410	.453
10	260	4.485	.405
11	254	4.203	.638
12	258	4.203	.509
13	259	4.472	.451
14	259	4.491	.428
15	259	4.416	.438
16	259	4.555	.417
17	259	4.222	.608
18	256	4.431	.506
19	259	4.250	.529
20	259	4.587	.399
21	259	4.392	.470
22	256	4.382	.597
23	259	4.318	.483
24	259	4.279	.530
25	259	4.011	.619
26	259	4.046	.669
27	259	4.321	.512
28	259	4.389	.462
29	259	4.643	.313
30	259	4.430	.474
31	259	3.969	.543
32	259	4.012	.668
33	259	4.676	.324
34	259	4.160	.512
35	260	3.985	.614
36	260	4.095	.624

alpha = .97

Table 2
T-tests for Undergraduate/Graduate Mean Scores on Items 9-36

Item	Mean		t	p
	Undergraduate	Graduate		
9	4.30	4.42	1.38	.180
10	4.53	4.54	.64	.529
11	4.15	4.02	-.68	.502
12	4.00	4.27	2.00	.057
13	4.25	4.49	2.24	.034
14	4.48	4.50	.34	.739
15	4.40	4.39	-.19	.849
16	4.51	4.57	.61	.550
17	4.10	4.21	.71	.484
18	4.30	4.50	1.89	.071
19	4.17	4.25	.67	.509
20	4.47	4.62	1.66	.110
21	4.34	4.44	.89	.380
22	4.22	4.45	2.06	.050
23	4.24	4.29	.47	.643
24	4.24	4.30	.40	.689
25	3.95	4.00	.37	.713
26	3.83	4.08	2.11	.045
27	4.28	4.34	.49	.627
28	4.35	4.38	.31	.756
29	4.64	4.57	-.78	.445
30	4.46	4.51	.57	.572
31	3.90	4.26	3.29	.003
32	3.86	3.99	1.15	.260
33	4.61	4.67	.69	.500
34	4.08	4.27	1.76	.091
35	3.80	4.02	1.68	.106
36	4.02	4.19	1.07	.295

Underlined p-values are statistically significant at the .05 level
N = 26

Table 3

Means and p value for t-test for Item2 Responses

Item	Response Means A (It was required)	B (Advisors recommendation)	C (Subject was of interest)	D (Teachers excellent reputation)	t-tests p C/A	B/A	(A + B)/(C + D)
9	4.35	4.46	4.51	4.68	.001	.040	.005
10	4.47	4.53	4.62	4.76	.001	.007	.001
11	4.09	4.28	4.40	4.47	.002	.440	.032
12	4.14	4.28	4.28	4.39	.033	.384	.120
13	4.40	4.54	4.46	4.65	.001	.145	.105
14	4.43	4.60	4.58	4.73	.001	.110	.002
15	4.41	4.43	4.56	4.66	.002	.009	.001
16	4.53	4.67	4.65	4.70	.029	.697	.136
17	4.31	4.25	4.31	4.63	.001	.001	.001
18	4.41	4.49	4.44	4.62	.030	.609	.547
19	4.24	4.40	4.33	4.50	.003	.291	.118
20	4.57	4.66	4.68	4.87	.001	.004	.001
21	4.33	4.45	4.45	4.64	.001	.040	.016
22	4.37	4.46	4.41	4.60	.011	.609	.294
23	4.32	4.37	4.38	4.57	.003	.046	.022
24	4.25	4.31	4.40	4.62	.001	.001	.001
25	3.99	4.00	4.11	4.24	.053	.105	.032
26	4.02	4.15	4.26	4.20	.119	.645	.036
27	4.31	4.43	4.30	4.54	.056	.463	.626
28	4.37	4.54	4.46	4.67	.001	.071	.047
29	4.63	4.68	4.70	4.82	.001	.031	.013
30	4.34	4.50	4.60	4.67	.001	.035	.001
31	3.92	4.11	4.07	4.14	.128	.795	.346
32	4.02	4.11	4.20	4.20	.150	.458	.063
33	4.64	4.73	4.72	4.84	.006	.109	.036
34	4.10	4.20	4.20	4.35	.007	.128	.070
35	3.97	4.20	4.20	4.42	.001	.072	.001
36	4.11	4.30	4.20	4.55	.001	.008	.011

Underlined p-values are significant at the .05 level
Adjustments for multiple t-tests are included

Table 4
Means, t value, and p values for Item6

Item	Mean Scores: Percent of Classes Attended 60 - 80%	80 - 100%	t	p
9	4.36	4.43	1.87	.063
10	4.44	4.51	2.06	.040
11	4.19	4.22	0.84	.405
12	4.17	4.23	1.17	.244
13	4.48	4.53	0.92	.357
14	4.44	4.50	1.48	.141
15	4.37	4.45	2.12	.036
16	4.54	4.56	0.62	.538
17	4.18	4.26	1.59	.114
18	4.39	4.47	2.18	.031
19	4.22	4.22	0.10	.905
20	4.53	4.58	1.63	.105
21	4.38	4.40	0.46	.646
22	4.35	4.40	1.18	.240
23	4.27	4.34	1.64	.102
24	4.25	4.28	0.60	.549
25	4.00	4.05	1.01	.314
26	3.96	4.08	2.20	.029
27	4.34	4.34	0.24	.802
28	4.36	4.39	0.75	.450
29	4.61	4.65	1.45	.149
30	4.35	4.46	2.54	.012
31	3.85	4.08	3.08	.002
32	3.97	4.03	1.09	.278
33	4.64	4.68	1.25	.213
34	4.13	4.20	1.87	.063
35	3.94	4.01	1.56	.120
36	4.04	4.08	0.93	.352

Underlined p-values are significant at the .05 level
N = 204- 216

Table 5
Means and p values of the t-tests for Item8

Item	Means Level of Interest			p values	
	Very Great	Average	Small	VG/A	VG/S
9	4.31	4.41	4.39	.247	.400
10	4.54	4.30	4.20	.001	.003
11	4.34	4.02	4.01	.005	.004
12	4.19	4.04	3.98	.051	.200
13	4.55	4.36	4.25	.004	.060
14	4.56	4.34	4.31	.001	.048
15	4.49	4.20	4.14	.001	.004
16	4.69	4.47	4.48	.001	.003
17	4.40	4.24	4.04	.004	.001
18	4.54	4.28	4.25	.001	.002
19	4.36	3.96	3.89	.001	.001
20	4.68	4.45	4.40	.001	.008
21	4.46	4.20	4.15	.001	.016
22	4.44	4.25	4.22	.063	.040
23	4.44	4.21	4.18	.001	.006
24	4.40	4.10	3.90	.001	.001
25	4.14	3.84	3.60	.001	.001
26	4.13	3.75	3.80	.001	.016
27	4.44	4.18	4.20	.001	.022
28	4.50	4.18	4.11	.001	.001
29	4.70	4.51	4.49	.003	.016
30	4.43	4.22	4.08	.005	.003
31	4.06	3.79	3.84	.002	.167
32	4.07	3.95	3.68	.101	.003
33	4.72	4.51	4.52	.004	.025
34	4.28	3.96	3.80	.001	.001
35	4.24	3.67	3.42	.001	.001
36	4.23	3.93	3.83	.001	.003

Underlined p-values are significant at the .05 level
Adjustments for multiple t-tests are included
N = 68-95

Table 6

Factor Loadings for Item9-Item34

Item	Factor 1	Factor 2	Factor 3	Factor 4
24	0.866	0.162	0.010	-0.100
10	0.862	0.150	-0.027	-0.055
14	0.849	0.079	-0.052	0.053
15	0.834	0.184	-0.044	-0.062
30	0.796	-0.001	0.013	0.101
28	0.703	0.307	0.031	-0.026
19	0.697	0.303	-0.057	0.081
9	0.692	0.119	0.118	-0.079
13	0.680	-0.250	0.168	0.149
34	0.670	-0.018	0.276	0.138
21	0.415	0.350	0.240	0.079
31	0.394	0.193	0.111	0.199
16	- 0.033	0.939	0.083	-0.059
23	0.054	0.780	0.111	0.033
33	0.167	0.708	0.071	0.029
27	0.106	0.694	0.216	-0.004
29	0.267	0.599	0.016	0.088
12	0.267	0.591	-0.061	0.177
20	0.492	0.493	-0.022	-0.004
22	0.119	0.041	0.839	-0.057
18	0.034	0.148	0.810	-0.033
11	-0.121	0.067	0.698	0.150
25	0.432	0.062	0.534	-0.011
17	0.057	0.427	0.422	0.001
32	0.051	-0.058	0.102	0.869
26	0.023	0.178	-0.052	0.840

Table 7
Student Characteristic by Factor t-tests

Item	Factor1			Factor2			Factor3			Factor4		
	Mean	Contrast	p	Mean	Contrast	p	Mean	Contrast	p	Mean	Conrtast	p
1 - A: Graduate	51.11			31.02			19.52	A/ B	.011	7.96		
B: Undergraduate	51.90			31.14			17.35			8.11		
2 - A: Required	50.84	(A+B)/		30.87	(A+B)/		18.80			7.94	A/C	.004
B : Advisor	52.51	(C+D)	.000	31.76	(C+D)	.006	19.23			8.19		
C: Interest	52.53			31.59			19.69			8.40		
D: Reputation	56.18	B/D	.000	33.29			20.84			8.62		
3 - A: <2.5-2.29	48.41			29.65			19.94			7.68		
B: 2.5-2.99	49.00			30.01			20.40			7.71		
C: 3.0-3.49	48.81			29.79			19.76			7.72		
D: 3.5-4.0	50.20			30.47			20.28			7.96		
6 - D: 60-80	48.05	D/E	.001	29.38	D/E	.002	19.17			7.78		
E: 80-100	51.21			30.96			18.75			7.94		
7 - A: 0-2	49.68			30.17	C/A	.004	17.26			8.06	E/B	.003
B: 3-5	50.99			31.10			18.38			7.99	E/A	.002
C: 6-8	50.73			31.14			18.60			7.56	C/B	.005
D: 9-11	51.04			31.03			18.52			7.66		
E:>12	49.74			30.25			17.01			7.26		
8 - A: Nill	48.13	E/D	.004	30.62	E/C	.001	17.65	E/A	.005	6.98	E/C	.004
B: Small	47.70	E/C	.001	29.74	E/B	.002	18.02	D/A	.008	7.36	E/B	.002
C: Average	49.02	E/B	.001	30.12	D/C	.001	18.12			7.67	E/A	.001
D: Substantial	51.07	E/A	.001	31.22	D/B	.001	18.29			7.88	D/A	.001
E: Very great	52.35	D/C	.001	31.59			18.73			8.11		
		D/B	.001									

AERA CONFERENCE 1997

TM026812

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

ERIC

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Student Evaluation of University Teaching & Structure and Relationship with student characteristics

Author(s): Fadia Nasser & David Glassman.

Corporate Source:

Publication Date:

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following two options and sign at the bottom of the page.

	The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2 documents	
☑ Check here **For Level 1 Release:** Permitting reproduction in microfiche (4" x 6" film) or other ERIC archival media (e.g., electronic or optical) *and* paper copy.	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	☐ Check here **For Level 2 Release:** Permitting reproduction in microfiche (4" x 6" film) or other ERIC archival media (e.g., electronic or optical), but *not* in paper copy.
	Level 1	Level 2	

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Sign here→ please

Signature: Fadia Nasser

Printed Name/Position/Title: Fadia Nasser / Doctoral student

Organization/Address: University of Georgia, Athens GA, 30602

Telephone: 706 542-4396

FAX: (706) 542-4240

E-Mail Address: fnassir@uga.cc.uga.edu

Date: March 28, 1997